PHONOLOGICAL AND LEXICAL MISMATCH DETECTION IN 30-MONTH-OLDS AND ADULTS MEASURED BY PUPILLOMETRY

Tom Fritzsche & Barbara Höhle

Department of Linguistics, University of Potsdam, Germany
tom.fritzsche@uni-potsdam.de, hoehle@uni-potsdam.de

ABSTRACT

This study shows for the first time that mispronunciation detection in 30-month-old children and adults can be measured using pupillometry. Compared to correctly pronounced words we found that mispronounced ones result in larger pupil dilations. For unrelated labels, which could either be a word or a non-word, we found different effects in children and adults. Children's pupillary responses for all unrelated labels were not different from those for correct labels whereas in adults we observed an increase in pupil size for unrelated words but not for non-words.

Taking pupillary responses as an indicator of processing costs, we argue that for children as well as for adults a phonologically deviant word requires more resources to activate its matching entry in the mental lexicon compared to correctly pronounced words. Other measures like looking or reaction times are unable to capture this dimension in such a direct way.

Keywords: phonological processing, mispronunciation detection, pupillometry, language acquisition

1. INTRODUCTION

Detecting a mismatch between an expected word and an acoustically deviant one depends on the nature and specificity of the mental representation of that word and the type of mismatch. Most studies investigating children's phonological representations [4, 12, 14, 16] rely on target looks in a preferential looking paradigm where an auditory label is played while two pictures (one depicting the target, the other an unrelated referent) are presented on a screen. Target looks usually decrease as the amount of deviance (phonetic/phonological change) in the mismatching word increases [12, 17]. This inverse relationship, though indicative of processing fine-grained phonological information, does not signal the depth of processing that is required in these tasks.

Pupillary responses are taken to be a general "measure of processing load" [3, p. 159] with larger dilations reflecting more processing effort [6, 7]. Similar to the EEG signal, pupil size can be measured continuously online, although response latencies are larger compared to the electrophysiological signal. All remote eye-tracking systems provide pupil size data. An advantage of pupillometry over EEG-measures is that it is not necessary to attach equipment to the participant's head which makes this method especially suited for the use with young children. Pupillometric effects have recently been reported for semantic mismatch detection in picture-word pairs by adults [8] and could be correlated with the amplitude of the N400 component – an index of semantic integration costs [10].

The current study asks whether it is possible to detect responses to auditory mismatches with pupillometry in 30-month-old children and adults. If this is the case, it might provide a useful tool for assessing acoustic and phonological processing abilities, particularly in young children.

By the age of two years, young children's phonological representations are detailed enough to detect segmental changes in familiar and newly learned words [4, 12, 14, 16, 1, 11, 13, 18]. Pupillometry as a technique has previously been employed successfully in young children. Already 3-month-old's pupils dilate in response to an infrequent syllable in a sequence of a frequently repeated one [5]. A the age of 2.5 years, a pupil size increase has been observed for bilingual children when seeing unrelated compared to matching images after having heard a word [9]. This increase was not present in monolingual children of the same age. Against this background the prerequisites are met for being able to measure pupillary responses in 30-month-old children in a mispronunciation detection study. Adults will be included as a reference group.

2. METHOD

In this study we probe the phonological representations of familiar words by activating a stored word form with an image before presenting an auditory label. Contrary to preferential looking studies, showing only one image avoids potentially interfering influences created by a second distractor image. Including unrelated labels allows us to determine whether any mismatch or just phonetically or phonologically similar ones are associated with higher processing costs.

2.1. Participants

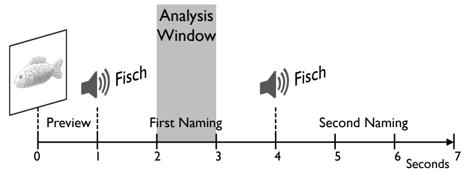
Two age groups were tested. 25 children (12 girls) with a mean age of 30.1 months (29.2–31.0) and 28 students (20 women) with a mean age of 24 years (18–38) entered the analysis. All were raised monolingually with German as first language and had no known visual or hearing deficits. Data of three additional children and one adult had to be removed from the analysis because not enough eye-tracking data (at least two trials per condition) could be acquired due to track loss, calibration error, or lacking compliance (one child). Parents were reimbursed for their participation, children received a small gift while students earned course credit.

2.2. Materials and design

Image and sound files were taken from a published study by Höhle and colleagues [4]. The ten images depicted referents of familiar words. The correct labels for these were words with a CV or CVC syllable structure, six of which served as experimental items and four as fillers. According to the German norms of the MacArthur-Bates Communicative Development Inventories [15], all words are produced by at least 60% of all 30-month-olds (mean: 89%). The images of the experimental items were paired with the sound files to yield three conditions: correct, mispronounced (mispro henceforth) and unrelated. Mispro labels were created by changing the place of articulation of the initial consonant such that the word resulted in a non-word (e.g. [fɪʃ] to [zɪʃ] for Fisch 'fish'). The unrelated condition was varied between participants: for half of them unrelated labels consisted of words, for the other half of non-words. We re-used correct labels from other images as words in the unrelated condition (e.g. [k^{h}uː] Kuh 'cow' in the example above). The non-words were derived by applying the mispro change to an unrelated word (e.g. [p^{h}uː]). Filler items always appeared with a correct label.

Image and sound presentation were synchronized (Fig. 1). The image was visible throughout the whole trial for seven seconds. After one second the label (in one of the three conditions) was played. The duration of the labels varied between 400 and 1200 msec (mean: 800 msec). Three seconds after the onset of the first naming the label was repeated. This was done as precaution in case children needed more time. The three conditions were tested in a within-participants design. Each experimental image appeared once in each condition (correct, mispro, and unrelated) while filler images were repeated three times resulting in 30 trials per participant.

Figure 1: Course of a trial. The presentation of the image continued throughout until the end.



The ratio of correct (including fillers) and mismatching (mispro, unrelated) trials was 3 to 2. Except for the filler items, labels were not repeated in different trials within a participant.

2.3. Procedure

Participants were tested individually in a dimly lit room without natural light to prevent changes in ambient luminance. After having obtained consent from the parents they sat in a lean-back chair with their child on the lap. During the test parents were asked to close their eyes. Adult participants filled out a questionnaire and a consent form before taking the seat in front of the eye-tracker. The screen was adjusted to be at a distance of 60 to 70 cm away from the participant's eyes. Following a five-point calibration with a moving circle on a dark background, the presentation of the test trials began. All participants were instructed to pay close attention to the pictures and the words. Every four trials an attention-getting stimulus (different animated cartoon characters) was presented for as long as the participant required or the experimenter deemed necessary. Upon finishing the study, parents and adults were debriefed and were given the opportunity to ask questions. For all experimental items it was checked with the parents that their child knew the correct label.

2.4. Apparatus and analysis

Eye movements and pupil size were recorded by a Tobii 1750 binocular corneal reflection eye-tracker with a temporal resolution of 50 Hz. Stimulus presentation and data acquisition were implemented in the ClearView software. According to the manufacturer's manual, spatial accuracy and recovery time after track loss are .5 to 1° and 100 msec, respectively. All stimuli appeared centrally on a 17" (1280 x 1024) TFT display with a size of 300 by 300 pixels subtending a horizontal and vertical viewing angle of 7.4°.

Only valid data entered the analysis, i.e. when at least one eye could be recorded correctly. Short intervals of missing data (up to 400 msec, usually the consequence of a blink) were linearly interpolated for each eye separately before combining both into one measure. Then, a baseline pupil dilation in the interval 500 msec prior to the onset of the first naming was calculated for each trial and subtracted from the following data points. The so obtained variation in pupil diameter is time-locked to the auditory presentation. Trials with more than 50% missing data or no data in the baseline were removed.

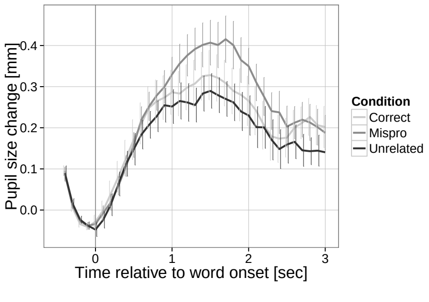
2.5. Predictions

Assuming that more processing costs are necessary for establishing a relation between a visually activated lexical entry and a mispronounced acoustic signal, we predict larger pupil dilations for the mispro vs. the correct condition. For unrelated labels, different scenarios are conceivable. Based on the findings by Kuipers and Thierry [8, 9], unrelated words compared to correct labels could trigger larger pupils in adults while for children (monolinguals in our sample) no difference might be found. For the condition of unrelated non-words we do not have an informed hypothesis. This manipulation was included to assess the influence of the lexical status of mismatching labels.

3. RESULTS

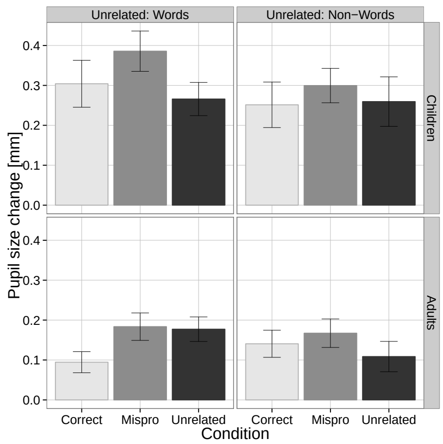
For children, 82% of all experimental trials entered the analysis, for adults 94%. Figure 2 illustrates the response dynamics of the pupil by depicting the first four seconds of all trials for those children who heard words as unrelated labels.

Figure 2: Time course of the pupil size changes for children in the unrelated word group.



The pupil size increases in response to the acoustic information in all three conditions. The mean peak latency measured from word onset is 1535 msec for children and 1562 msec for adults (averaged across all conditions). Thus, the analysis window from one to two seconds after word onset is justified by the data. The differences between the mean pupil sizes in the selected time window were analysed using a linear mixed-effects model [2] with the factors *Condition*, *Age*, and *Word Status* of the unrelated label. Correct items served as baseline against which the other other conditions are compared. The factor *Age* was specified such that all effects are tested for children alone and potential differences to adults show up as significant interactions with this factor. We included three variance components: for participants, for images, and for condition, that is we allowed the effect of *Condition* to vary across participants. Figure 3 shows the plotted grand averages aggregated over participants.

Figure 3: Mean pupil size changes in a window 1 to 2 sec after auditory word onset separated by age group and word status of the unrelated label. Error bars denote one standard error.



For the factor *Condition*, pupil changes to mispro labels are larger than to correct ones (coef.: .06, t = 2.23, p < .05) in both age groups by virtue of a missing interaction with *Age* (coef.: –.006, t = .156, p = .88). In children, responses to correct and unrelated labels were not different from each other (coef.: –.02, t = .663, p = .51) and *Word Status* had no effect (coef.: .05, t = .810, p = .42). An effect of *Age* revealed that adults' pupil size changes were overall smaller than children's (coef.: –.16, t = 3.64, p < .001). In addition, there is a marginally significant three-way interaction between *Age*, the unrelated *Condition*, and *Word Status* (coef.: .14, t = 1.78, p = .081) revealing that unlike in children there is an increase in adults' pupil dilations to unrelated *words* compared to correct labels while there is no such increase for unrelated non-words. All other interactions did not reach significance (all t-values < 1.18, all p-values > .24).

4. DISCUSSION

Consistent with our prediction, larger pupil dilations were observed for mispronounced labels compared to correct ones. The size of this effect is virtually identical in children and adults. This is in line with research showing that the phonological representations in young children include (sub-) segmental details. For unrelated labels, children do not exhibit increased pupil sizes compared to correct labels and the lexical status (word or non-word) of unrelated labels has no effect. It seems that children consider an unrelated label, be it a word or not, as not bearing a relation to the picture and thereby prevent any additional processing, which is only found in the case of mispronunciations. This relates to findings from Kuipers and Thierry [9] who have shown that bilingual but not monolingual 30-month-olds show larger pupil dilations in response to unrelated words. They argue that bilinguals "are more tolerant to variation in word-referent mappings" [9, p. 2853]. A similar explanation could hold for the adults in the current study. Even though they surely have noticed the mismatch between unrelated words and the picture, they might have tried to establish some connection between the two, which increased the processing load compared to correct labels. Regarding the non-words, adults do not seem to establish a relation between them and the pictures – just like the children.

An alternative interpretation for the marginal effect of the lexical status in adults – though speculative at this point – could be that the presence of non-words in the experiment changes the processing mode or strategy of the participants. In the absence of non-words (lower left panel in Fig. 3) mispronounced words (phonological mismatches) pattern with unrelated words (semantic mismatches). If non-words are encountered (lower right panel in Fig. 3) correct and mispronounced labels seem to pattern together so that it is not match vs. mismatch but rather lexical vs. non-lexical items.

Pupillometric responses to deviant acoustic labels are selective and not just a general marker of mismatch detection. It is the degree (low vs. high similarity to the correct label for mispro and unrelated conditions) and/or the type (phonological or semantic) of the mismatch that results in different amplitudes of the pupil response.

5. SUMMARY AND CONCLUSIONS

The present study compared correct and mismatching auditory labels to visually depicted referents. Results show that a mispronunciation in the initial consonant resulted in an increase in pupil size compared to correct labels and showed for the first time that mispronunciation detection can be measured with pupillometry. The fact that the pupil response to mismatching labels is modulated by the phonological overlap with the correct form makes this measurement especially interesting for the study of phonological development. It could be used in the future to explore whether the processing costs associated with detecting acoustic deviations are changing in a gradual or categorical manner in response to the signal and whether they reflect rather phonetic and/or phonological processes.

Differences between children and adults in response to unrelated words are suggestive of different resource allocations in the two age groups. While children seem to dismiss any relation between an unrelated word and a picture, adults seem to explore potential relationships.

Pupillometry, in our eyes, proves to be a valuable and promising technique for investigating online phonetic, phonological, and lexical processing, especially in young children. It can be taken as a window on resource allocation during language processing and serves as a direct measure of processing costs.

Acknowledgements: We would like to thank the children and their parents as well as the students for participating, Anna Richert for conducting the study, and the BabyLAB assistants Anne Beyer, Stefanie Meister, and Katja Schneller for arranging the appointments.

6. REFERENCES

[1] Ballem, K. D., Plunkett, K. 2005. Phonological specificity in children at 1;2. *Journal of Child Language* 32(1), 159–173.

[2] Bates, D. M., Maechler, M., Bolker, B., Walker, S. 2014. lme4: Linear mixed-effects models using Eigen and S4. http://CRAN.R-project.org/package=lme4. visited 5-Jun-14.

[3] Beatty, J., Lucero-Wagoner, B. 2000. *The pupillary system* 142–162. Cambridge, UK: Cambridge University Press 2nd edition.

[4] Höhle, B., van de Vijver, R., Weissenborn, J. 2006. Word processing at 19 months at its relation to language performance at 30 months: A retrospective analysis of data from german learning children. *Advances in Speech and Language Pathology* 8(4),

356–363.

[5] Hochmann, J.-R., Papeo, L. 2014. What infants' pupils say about speech: The invariance problem in 3- and 6-month-old infants. *Psychological Science* 25(11), 2038–2046.

[6] Just, M. A., Carpenter, P. A. 1993. The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology* 47(2), 310–339.

[7] Kahneman, D., Beatty, J. 1966. Pupil diameter and load on memory. *Science* 154(3756), 583–1585.

[8] Kuipers, J. R., Thierry, G. 2011. N400 amplitude reduction correlates with an increase in pupil size. *Frontiers in Human Neuroscience* 5(61).

[9] Kuipers, J.-R., Thierry, G. 2013. Erp-pupil size correlations reveal how bilingualism enhances cognitive flexibility. *Cortex* 49(10), 2853–2860.

[10] Kutas, M., Hillyard, S. A. 1980. Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology* 11(2), 99–116.

[11] Mani, N., Mills, D. L., Plunkett, K. 2012. Vowels in early words: An event-related potential study. *Developmental Science* 15(1), 2–11.

[12] Mani, N., Plunkett, K. 2011. Does size matter? subsegmental cues to vowel mispronunciation detection. *Journal of Child Language* 38(3), 606 –627.

[13] Swingley, D. 2003. Phonetic detail in the developing lexicon. *Language and Speech* 46(2-3), 265–294.

[14] Swingley, D., Aslin, R. N. 2000. Spoken word recognition and lexical representation in very young children. *Cognition* 76(2), 147–166.

[15] Szagun, G., Stumper, B., Schramm, S. A. 2009. *Fragebogen zur frühkindlichen Sprachentwicklung (FRAKIS) und FRAKIS-K (Kurzform)*. Frankfurt: Pearson Assessment.

[16] White, K. S., Morgan, J. L. 2008. Sub-segmental detail in early lexical representations. *Journal of Memory and Language* 59(1), 114–132.

[17] White, K. S., Yee, E., Blumstein, S. E., Morgan, J. L. 2013. Adults show less sensitivity to phonetic detail in unfamiliar words, too. *Journal of Memory and Language* 86(4), 362–378.

[18] Yoshida, K. A., Fennell, C. T., Swingley, D., Werker, J. F. 2009. Fourteen-month-old infants learn similar-sounding words. *Developmental Science* 12(3), 412–418.